



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER OF PATENTS AND TRADEMARKS  
Washington, D.C. 20231  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/576,116	05/22/2000	Matthew Lennig	003932.P014	2810

7590 03/17/2003

Jordan M. Becker  
Blakely, Sokoloff, Taylor, & Zafman LLP  
12400 Wilshire Boulevard 7th Floor  
Los Angeles, CA 90025

EXAMINER

AZAD, ABUL K

ART UNIT

PAPER NUMBER

2654

DATE MAILED: 03/17/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

**Office Action Summary**

Application No.

09/576,116

Applicant(s)

LENNIG, MATTHEW

Examiner

ABUL K. AZAD

Art Unit

2654

v

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 22 May 2000.
- 2a) ☐ This action is **FINAL**.                      2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-41 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☒ Claim(s) 20 and 41 is/are allowed.
- 6) ☐ Claim(s) 1-15, 17, 18 and 21-40 is/are rejected.
- 7) ☒ Claim(s) 16 and 19 is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 22 May 2000 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- 11) ☐ The proposed drawing correction filed on \_\_\_\_\_ is: a) ☐ approved b) ☐ disapproved by the Examiner.
- If approved, corrected drawings are required in reply to this Office action.
- 12) ☐ The oath or declaration is objected to by the Examiner.

**Priority under 35 U.S.C. §§ 119 and 120**

- 13) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.
- 14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).
- a) ☐ The translation of the foreign language provisional application has been received.
- 15) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

**Attachment(s)**

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) 4.
- 4) ☐ Interview Summary (PTO-413) Paper No(s). \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: \_\_\_\_\_

**DETAILED ACTION**

1. Claims 1-41 are pending in this action.

***Information Disclosure Statement***

2. The reference cited in the information Disclosure statement IDS, PTO-1449, Paper No. 4 has been considered.

***Claim Objections***

3. Claims 3 and 4 are objected to under 37 CFR 1.75(c), as being of improper dependent form for failing to further limit the subject matter of a previous claim. Applicant is required to cancel the claim(s), or amend the claim(s) to place the claim(s) in proper dependent form, or rewrite the claim(s) in independent form. Claims 3 and 4 are depend on each other, but they are failing to further limit the subject matter of a previous claim.

***Claim Rejections - 35 USC § 102***

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Art Unit: 2654

5. Claims 1-4, 6-7, 9-11, 14-15, 17-18, 21-30, 32-38 and 40 are rejected under 35 U.S.C. 102(b) as being anticipated by Peckham et al. (EP 0 424 071).

As per claim 1, Peckham teaches, "a method comprising":

"inputting speech representing an utterance and having an intonation" (Page 5, lines 32-33, particularly reads on "the input words are analysed to extract normalized cepstral coefficients and pitch" where "intonation" reads on "pitch"); and

"identifying an endpoint of the utterance based on the intonation" (Page 14, lines 55-56, particularly reads on "the use of pitch information, preferably in combination with energy, in identifying the start and end points of utterances", where "intonation" reads on "pitch").

As per claim 2, the claim limitation is rejected based on the rational given to claim 1 above, and further Peckham teaches, "wherein said identifying an endpoint of the utterance based on the intonation comprises comparing the intonation with an intonation model" (Page 14, lines 36-56, particularly reads on "such as pitch and delta cepstrum may be used in the enrolment and verification process").

As per claim 3, Peckham teaches, "further comprising determining the intonation by computing the fundamental frequency of the utterance" (Page 5, lines 32-33, particularly reads on "the input words are analysed to extract normalized cepstral coefficients and pitch" where pitch by definition is the fundamental frequency, see text book of Deller et al.).

As per claim 4, Peckham teaches, "wherein said determining the intonation comprises using an intonation model to determine the intonation" (Page 14, lines 36-56,

Art Unit: 2654

particularly reads on "such as pitch and delta cepstrum may be used in the enrolment and verification process").

As per claim 6, the claim limitation is rejected based on the rational given to claim 1 above, and further Peckham teaches, "determining a period of time that has elapsed since the speech dropped below a threshold value" (Page 8, lines 50-56, particularly reads on "this system looks backwards in time from the beginning of the period and forwards in time from the end of this period to discover the points where the energy falls to 10 per cent of the maximum values. This points are used to identify the start and end of the spoken word for analysis"); and

"wherein said identifying an endpoint of the utterance comprises identifying the endpoint of the utterance further based on the period of time" (Page 8, lines 50-56, particularly reads on "this system looks backwards in time from the beginning of the period and forwards in time from the end of this period to discover the points where the energy falls to 10 per cent of the maximum values. This points are used to identify the start and end of the spoken word for analysis").

As per claim 7, the claim limitation is rejected based on the rational given to claim 1 above, and further Peckham teaches, "wherein said identifying an endpoint of the utterance comprises identifying the endpoint of the utterance further based on a length of time for which an energy value of the speech has remained below a predetermined energy value" (Page 8, lines 50-56, particularly reads on "this system looks backwards in time from the beginning of the period and forwards in time from the end of this period

to discover the points where the energy falls to 10 per cent of the maximum values. This points are used to identify the start and end of the spoken word for analysis”).

As per claims 25-30, 32-38 and 40, they are interpreted and thus rejected for the same reasons set forth in the rejection of claims 1-4, 6-7, because essentially they have similar limitations and scope.

As per claim 9, Peckham teaches, “a method of operating an endpoint detector”, the method comprising:

“inputting speech representing an utterance, the utterance having an intonation” (Page 5, lines 32-33, particularly reads on “the input words are analysed to extract normalized cepstral coefficients and pitch” where “intonation” reads on “pitch”); and

“comparing the intonation of the utterance with an intonation model” (Page 14, lines 36-56, particularly reads on “such as pitch and delta cepstrum may be used in the enrolment and verification process”);

“determining a probability based on a result of said comparing” (Page 14, line 36 to Page 15, line 11, particularly reads on “preferably in combination with the use of a statistical method of combining the computed similarity values into single measure”, where statistical method according to Page 5, lines 7-8 is trial and error “probability”); and

“identifying an endpoint of the utterance based on the probability” (Page 14, line 36 to Page 15, line 11, particularly reads on “preferably in combination with the use of a statistical method of combining the computed similarity values into single measure”).

As per claim 10, the claim limitation is rejected based on the rational given to claim 9 above, and further Peckham teaches, "further comprising determining the intonation of the utterance as a function of the fundamental frequency of the utterance" (Page 5, lines 32-33, particularly reads on "the input words are analysed to extract normalized cepstral coefficients and pitch" where pitch by definition is the fundamental frequency, see text book of Deller et al.).

As per claim 11, the claim limitation is rejected based on the rational given to claim 9 above, and further Peckham teaches, "determining a period of time that has elapsed since a value of the speech dropped below a threshold value" (Page 8, lines 50-56, particularly reads on "this system looks backwards in time from the beginning of the period and forwards in time from the end of this period to discover the points where the energy falls to 10 per cent of the maximum values. This points are used to identify the start and end of the spoken word for analysis"); and

"wherein said identifying an endpoint of the utterance comprises identifying the endpoint of the utterance further based on the period of time" (Page 8, lines 50-56, particularly reads on "this system looks backwards in time from the beginning of the period and forwards in time from the end of this period to discover the points where the energy falls to 10 per cent of the maximum values. This points are used to identify the start and end of the spoken word for analysis").

As per claim 14, Peckham teaches, "a method of operating an endpoint detector for speech recognition", the method comprising:

Art Unit: 2654

"inputting speech representing an utterance" (Page 5, lines 32-33, particularly reads on "the input words are analysed to extract normalized cepstral coefficients and pitch" where "intonation" reads on "pitch");

"determining that a value of the speech has dropped below a threshold value" (Page 8, lines 50-56, particularly reads on "this system looks backwards in time from the beginning of the period and forwards in time from the end of this period to discover the points where the energy falls to 10 per cent of the maximum values. This points are used to identify the start and end of the spoken word for analysis");

"computing an intonation of the utterance" (Page 5, lines 32-33, particularly reads on "the input words are analysed to extract normalized cepstral coefficients and pitch" where "intonation" reads on "pitch");

"referencing the intonation of the utterance against an intonation model to determine a first end-of-utterance probability" (Page 14, line 36 to Page 15, line 11, particularly reads on "preferably in combination with the use of a statistical method of combining the computed similarity values into single measure", where statistical method according to Page 5, lines 7-8 is trial and error "probability");

"determining a period of time that has elapsed since the value of the speech dropped below the threshold value" (Page 8, lines 50-56, particularly reads on "this system looks backwards in time from the beginning of the period and forwards in time from the end of this period to discover the points where the energy falls to 10 per cent of the maximum values. This points are used to identify the start and end of the spoken word for analysis");



"referencing the period of time against an elapsed time model to determine a second end-of-utterance probability" (Page 14, line 36 to Page 15, line 11, particularly reads on "preferably in combination with the use of a statistical method of combining the computed similarity values into single measure", where statistical method according to Page 5, lines 7-8 is trial and error "probability");

"computing an overall end-of-utterance probability as a function of the first and second end-of-utterance probabilities" (Page 14, line 36 to Page 15, line 11, particularly reads on "preferably in combination with the use of a statistical method of combining the computed similarity values into single measure", where statistical method according to Page 5, lines 7-8 is trial and error "probability"); and

"determining whether an end-of-utterance has occurred based on the overall end-of-utterance probability" (Page 14, line 36 to Page 15, line 11, particularly reads on "preferably in combination with the use of a statistical method of combining the computed similarity values into single measure", where statistical method according to Page 5, lines 7-8 is trial and error "probability").

As per claim 15, the claim limitation is rejected based on the rationale given to claim 14 above, and further Peckham teaches, "wherein said computing an intonation of the utterance comprises computing an intonation of the utterance by determining the fundamental frequency of the utterance as a function of time" Page 5, lines 32-33, particularly reads on "the input words are analysed to extract normalized cepstral coefficients and pitch" where pitch by definition is the fundamental frequency, see text book of Deller et al.).

As per claims 17 and 18, they are interpreted and thus rejected for the same reasons set forth in the rejection of claims 14 and 15, because claims are directed towards a method of operating an endpoint detector for speech recognition, essentially have similar limitations and scope.

As per claim 21, Peckham teaches, "a method of operating an endpoint detector for speech recognition", the method comprising:

"inputting speech representing an utterance" Page 5, lines 32-33, particularly reads on "the input words are analysed to extract normalized cepstral coefficients and pitch" where "intonation" reads on "pitch");

"determining an intonation of the utterance" Page 5, lines 32-33, particularly reads on "the input words are analysed to extract normalized cepstral coefficients and pitch" where "intonation" reads on "pitch");

"if the intonation of the utterance is determined to be generally decreasing, then setting a threshold time period equal to a first time value" (Page 6, line 39 to Page 7, line 43, particularly reads on "it is also important that this period is chosen to be greater than the pitch period of the lowest pitch frequency which the pitch analyzer is able to recognize", where "first time value" is "15 ms");

"if the intonation of the utterance is determined not to be generally decreasing, then setting the threshold time period equal to a second time value larger than the first time value" Page 6, line 39 to Page 7, line 43, particularly reads on "it is also important that this period is chosen to be greater than the pitch period of the lowest pitch

Art Unit: 2654

frequency which the pitch analyzer is able to recognize", where "second time value" is "25 ms"); and

"identifying an endpoint of the utterance based on the threshold time period" (Page 8, lines 29-56, teaches an endpoint is determined based on the threshold time period).

As per claim 22, the claim limitation is rejected based on the rational given to claim 21 above, and further Peckham teaches, "wherein said using the threshold time period to identify an endpoint of the utterance comprises using the threshold time period to identify an endpoint of the utterance by determining that an endpoint of the utterance has occurred if an energy value of the speech remains below a predetermined value for the threshold time period" Page 8, lines 50-56, particularly reads on "this system looks backwards in time from the beginning of the period and forwards in time from the end of this period to discover the points where the energy falls to 10 per cent of the maximum values. This points are used to identify the start and end of the spoken word for analysis").

As per claim 23, the claim limitation is rejected based on the rational given to claim 21 above, and further Peckham teaches, "wherein said determining an intonation of the utterance comprises using an intonation model" Page 14, lines 36-56, particularly reads on "such as pitch and delta cepstrum may be used in the enrolment and verification process").

As per claim 24, it is interpreted and thus rejected for the same reasons set forth in the rejection of claims 21-23.

***Claim Rejections - 35 USC § 103***

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. Claim 5 is rejected under 35 U.S.C. 103(a) as being unpatentable over Peckham et al. (EP 0 424 071) as applied to claim 1 above, and further in view of Zhao et al. (US 6,480,823).

As per claim 5, the claim limitation is rejected based on the rationale given to claim 1 above, further, Peckham teaches, "wherein said identifying the endpoint of the utterance comprises identifying the endpoint of the utterance based on a plurality of knowledge sources, wherein one of the knowledge sources is intonation" (Page 8, lines 29-56, where plurality of knowledge sources are pitch (intonation), energy and time etc.). Peckham does not teach referencing the input speech against a histogram based on training data for each of the knowledge sources. However, Zhao teaches, a histogram database (Fig. 1, elements 38 and 40). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to build a histogram database for each of the knowledge source because Zhao teaches the invention will detect both the beginning and end of speech as well as handling situations where the beginning of speech may have been lost through truncation will provide a better detection of speech in the noise condition (col. 1, lines 54-58).

8. Claims 8, 12, 13, 31 and 39 are rejected under 35 U.S.C. 103(a) as being unpatentable over Peckham et al. (EP 0 424 071) as applied to claims 7, 9, 25 and 33 above, and further in view of Megata et al. (JP 403245700).

As per claim 8, the claim limitation is rejected based on the rational given to claim 7 above, and further, Peakman does not explicitly teach identifying the endpoint of the utterance based on the duration of the final syllable of the utterance. However, Megata teaches, identifying the endpoint of the utterance based on the duration of the final syllable of the utterance (see purpose and constitution). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to identify an endpoint of a continues speech using the final syllable of the utterance as taught by Megata because an skilled artisan would readily recognized that would particularly detect the end point of the utterance, which helps enhancement of the recognition process.

As per claim 12, 31 and 39, they are interpreted and thus rejected for the same reasons set forth in the rejection of claim 8, because claim 12, 31 and 39 essentially have similar limitation and scope.

As per claim 13, the claim limitation is rejected based on the rational given to claim 12 above, and further Peckham teaches, "wherein said identifying an endpoint of the utterance comprises identifying the endpoint of the utterance further based on a period of time for which an energy value of the speech has remained below a threshold value" (Page 8, lines 50-56, particularly reads on "this system looks backwards in time from the beginning of the period and forwards in time from the end of this period to

Art Unit: 2654

discover the points where the energy falls to 10 per cent of the maximum values. This points are used to identify the start and end of the spoken word for analysis").

***Allowable Subject Matter***

9. Claims 20 and 41 are allowed.
10. Claims 16 and 19 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.
11. The following is a statement of reasons for the indication of allowable subject matter: the prior art of record fails to teach or fairly suggest "computing an overall end-of-utterance probability comprises computing the overall end-of-utterance probability as a function of the first, second, and third end-of-utterance probabilities".

***Conclusion***

12. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Mizuno et al., (US 5,732,392) teach method for speech detection in a high-noise environment.

Chow et al., (US 5,692,104) teach method and apparatus for detecting end points of speech activity.

**Contact Information**

13. Any inquiry concerning this communication or earlier communications from the examiner should be directed to **Abul K. Azad** whose telephone number is **(703) 305-3838**.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, **Marsha D. Banks-Harold**, can be reached at **(703) 305-4379**.

Any response to this action should be mailed to:

**Commissioner for Patents**

**Washington, D.C. 20231**

Or faxed to:

**(703) 872-9314**

(For informal or draft communications, please label "PROPOSED" or "DRAFT")

Hand-delivered responses should be brought to Crystal Park II, 2121 Crystal Drive, Arlington, VA, Sixth Floor (Receptionist).

Any inquiry of a general nature or relating to the status of this application should be directed to the Technology Center's Customer Service Office whose telephone number is **(703) 306-0377**.

  
Abul K. Azad

March 10, 2003